

Framework for Incorporating Downscaled Climate Output into Existing Engineering Methods: Application to Precipitation Frequency Curves

Accepted Pre-Print, ASCE Journal of Infrastructure Systems

Lauren M. Cook¹, S.M.ASCE, Christopher J. Anderson², and Constantine Samaras³, A.M.ASCE

Abstract: To improve the resiliency of designs, particularly for long-lived infrastructure, current engineering practice must be updated to incorporate a range of future climate conditions that are likely to be different from the past. However, a considerable mismatch exists between climate model outputs and the data inputs needed for engineering designs. The present work provides a framework for incorporating climate trends into design standards and applications, including: selecting the appropriate climate model source based on the intended application, understanding model performance and uncertainties, addressing differences in temporal and spatial scales, and interpreting results for engineering design. The framework is illustrated through an application to depth-duration-frequency curves, which are commonly used in stormwater design. A change factor method is used to update the curves in a case study of Pittsburgh, PA. Extreme precipitation depth is expected to increase in the future for Pittsburgh for all return periods and durations examined, requiring revised standards and designs. Doubling the return period and using historical, stationary values may enable adequate design for short duration storms; however, this method is shown to be insufficient to enable protective designs for larger duration storms.

Keywords: downscaled climate models, uncertainty, depth-duration-frequency, stormwater design

Introduction

Record-breaking rainfall has triggered more than 20 severe flood events in parts of Texas, Oklahoma, Louisiana, Arkansas, Missouri, Iowa, Florida, North Carolina, and South Carolina in 2015 and 2016. These events have led to the closure of two airports, flooding of more than 200 homes, numerous evacuations, cars stalled in high water requiring rescue, and deadly flash flooding. High water also led to spillway activation to protect New Orleans, as well as structural failure of more than 100 roads and retaining walls (Erdman 2016). Existing infrastructure systems were inadequate to deal with these events, which occurred outside of historical experience frequency. Design standards rely on historical observations and the assumption that climate is stationary (i.e., climate will not change over time). However, recent events and

numerous simulations of future climate conditions indicate that the past is no longer a reliable indicator of the conditions under which infrastructure will have to perform in the future (Walsh et al. 2014; Milly et al. 2008).

Climate change has the potential to affect infrastructure systems in multiple ways, including: (i) changes in average and/or extreme temperatures; (ii) variations in frequencies, intensities, and duration of precipitation causing extreme rainfall and flooding in some regions; (iii) changes in storm tracks and severe weather; (iv) an increase in sea levels and the risk of storm surge; and (v) a decrease of water availability in some areas (Walsh et al. 2014; IPCC 2014; Kilgore et al. 2016).

Recently, increased attention has been directed to infrastructure reliability (the ability of systems to remain functional during a disaster) and

¹ Graduate Research Assistant, Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA, United States

² Research Assistant Professor, Department of Agronomy, Iowa State University, Ames, Iowa

³ Assistant Professor, Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA, United States. Corresponding Author: csamaras@cmu.edu

resiliency (the ability to resist, absorb, and adapt to disruptions) (Faturechi and Miller-Hooks 2014). To ensure reliable and resilient infrastructure, engineering design standards must account for anticipated future conditions (Milly et al. 2008; Olsen 2015; Mailhot and Duchesne 2009; Moss et al. 2013; Barros and Evans 1997). These standards are set by organizations such as the American Society of Civil Engineers (ASCE) (ASCE 2013), agencies such as the Federal Highway Administration (FHWA) (Kilgore et al. 2016) or the National Oceanic and Atmospheric Administration (NOAA) (Bonnin et al. 2006), or by collaborations among organizations (e.g., 10 States Standards (Wastewater Committee of the Great Lakes - Upper Mississippi River 2014)).

One of the most advanced tools available to decision makers seeking to increase reliability and resilience of infrastructure is the use of high-resolution, or “downscaled,” climate models. Compared to general circulation models (GCMs) that simulate global climate systems, these downscaled models provide insight into localized conditions by generating finer-scale (4 – 50 km), future projections of air temperature, precipitation, evapotranspiration, wind speed, and other factors that affect regional patterns (Hall 2014). Most models agree on the direction of temperature change; however, for precipitation there are variations in trend and magnitude across models and geographic regions, leading to large uncertainty in results. For precipitation data particularly there is often a mismatch in the spatial and temporal resolution of the downscaled climate model and the micro scale (e.g., < 1 km) of inputs needed for engineering design standards and applications. Furthermore, the use of climate models introduces uncertainties and complicates data extraction and preparation requirements, compared to the current use of recorded historical data. A clear path from climate model predictions to development of updated design standards is needed.

Despite these challenges, by building on historical observations, scientists have successfully used global and downscaled climate models to inform higher spatial and temporal resolution precipitation trends for engineering applications. Weather generators, which can be adapted to different anticipated changes in climate, have been used to simulate synthetic, rainfall time series at the

station (point) scale at monthly, daily and hourly time steps (Wilks and Wilby 1999; Kilsby et al. 2007; Willems et al. 2013). Quantile mapping has been used to apply expected changes to the empirical distribution of observed rainfall events at the temporal and spatial resolution of the observations (Laflamme et al. 2016; Boé et al. 2007; Gudmundsson et al. 2012; Wood et al. 2004). Numerous studies utilize a “delta” or “change factor” technique, which applies the expected absolute (delta) or relative (ratio) change between current and future gridded projections to historical rainfall data (Wilks and Wilby 1999; Boé et al. 2007; Wood et al. 2004; K. Arnbjerg-Nielsen et al. 2013; Forsee and Ahmad 2011). These climate-informed local-scale models have also been used to update intensity-duration-frequency curves used in design of infrastructure affected by rainfall (Chandra et al. 2015; Cheng and AghaKouchak 2014; Forsee and Ahmad 2011; Zhu 2012; Kuo et al. 2015; Hassanzadeh et al. 2013; Mirhosseini et al. 2013).

Applications of these climate-informed methods can provide important insights; however, many reported studies provide insufficient detail regarding the importance and difficulty of obtaining a reliable historical record; selecting and extracting the appropriate climate output source; accounting for reliability and uncertainty in climate modeling; and incorporating findings into infrastructure planning and design. In the absence of a consensus on methods to update design standards to account for climate change, many stakeholders avoid the use of climate model output. Further, there is the potential for misuse through simplified choices, such as using output from a single climate model instead of an ensemble (or group) of models, or failing to account for model reliability and uncertainty in the interpretation of results. Given the widespread use of infrastructure design standards and the potential consequences to the public if they are improperly applied (including failure due to under-design or misallocation of taxpayer dollars due to overdesign), it is critical

that the most advanced and appropriate methods are used to update standards and that the challenges and limitations associated with this updating are well understood by those who will apply these techniques.

With this problem in mind, a five-step framework is proposed that can guide the revision of design standards, as well as engineering practice, through the use of publicly available, downscaled climate model outputs of future precipitation. By applying the framework, engineers will be able to define relevant aspects of the historical method that need to be updated; select the relevant climate data sources and extract output; manage model reliability and bound uncertainty; adjust for spatial and temporal resolution, and apply results to engineering design under climate non-stationarity. In the present work, as a demonstration, the framework is applied to a common input to stormwater infrastructure design: depth-duration-frequency (DDF) curves. These curves and their application will determine the performance and resiliency of stormwater infrastructure during future extreme events.

Framework Steps

The steps of the framework for updating engineering design standards are: (0) Define the existing design standard (or application) that relies on precipitation information; (1) Understand the historical requirements for existing standard and retrieve data; (2) Access appropriate climate model output based on requirements for the existing application; (3) Account for climate model uncertainty and reliability; (4) Incorporate climate model output into the required engineering format; and (5) Interpret results and incorporate changes into design practice. A flow chart of these steps is presented in Figure 1. Solid arrows display the suggested sequence of the steps from 0 to 5; dashed arrows represent the flow of information or data from Step 1 to Step 4.

Step 0. Define the existing design standard (or application) that relies on precipitation information

Standards for engineering design have been developed for a variety of engineering applications that are expected to be affected by a non-stationary climate, including: water supply management,

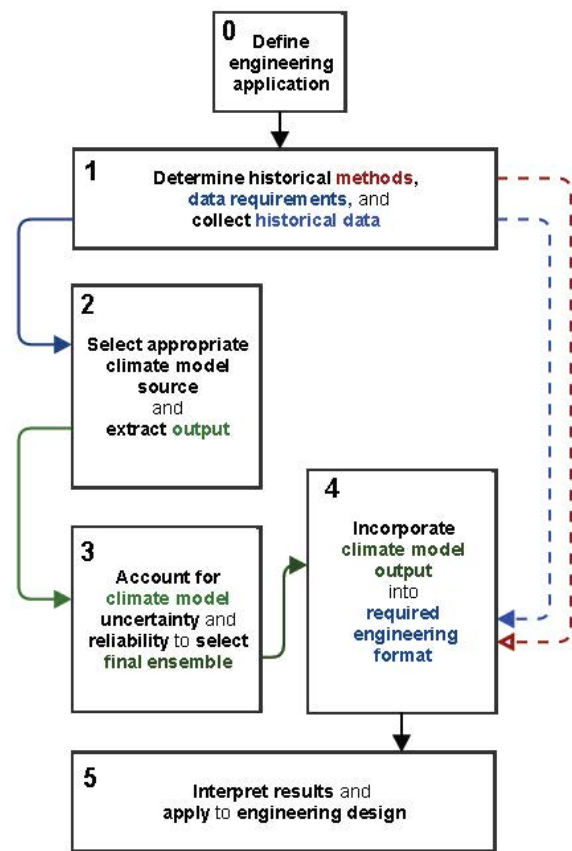


Fig 1. Framework for incorporating downscaled climate data into existing engineering applications (Note: solid arrows display the suggested sequence of the steps from 0 to 5; dashed arrows represent the flow of information or data from Step 1 to Step 4)

water quality regulations, flood forecasting, stormwater management, and wastewater collection and treatment. These and many other applications rely on different types of estimates of expected precipitation for a region. For example, floodplain delineation and stormwater management rely on duration-specific estimates of rainfall depth from intensity-duration-frequency (IDF) curves; whereas wastewater collection and treatment system design requires a peaking factor, usually relating to maximum daily or monthly rainfall.

Some applications may require a time series for precipitation (a sequence of data points collected over a time period, usually provided at evenly spaced time intervals). These different specific precipitation data determine the type of modifications that will be needed, and thus, defining the ways that rainfall data are used in the current standard is the initial step to updating.

Step 1. Understand historical requirements for existing standard and retrieve data

In this step, the engineer defines the nature of the data on which the design standard was based and obtains the historical data to enable re-creation of the supporting calculations underlying the current standard. Definition of the data includes the length of record, as well as the temporal and spatial resolution of the data required as input to re-create the components of the method. These specifications are important as they dictate the source of climate output that will be needed in step 2. Table 1 provides information on spatial and temporal resolutions required for analysis of several types of engineering applications and design standards. Large-scale optimization models used for reservoir or drought management use monthly or seasonal data, while stream flow and water quality simulations require daily or hourly data. Continuous hydrologic simulation models (e.g., Environmental Protection Agency (EPA) Storm Water Management Model (SWMM) (USEPA 2015)) use data at a sub-hourly time step and 1 km spatial resolution (Wood et al. 2000; Wilks and Wilby 1999), while IDF curves use multi-decadal time series of observed rainfall at individual geographic locations (point measurements) for durations ranging from 5 minutes to 72 hours (CSA 2012; Bonnin et al. 2006).

In support of these different data needs, historical precipitation data, collected through rain gauges, can be obtained at the point or grid scale. Many regional airports and local stormwater agencies collect rain gauge data at specific locations (points), at hourly intervals or less. Airport records tend to be longest (50 to 100 years); however, local agency data may be available at higher resolution (sub-hourly, multiple gauge sites) for a shorter time period. Another paper suggested that the usefulness of a rain gauge network is dependent on the density of gauges, the number of years of data, the type of rain gauge, and frequency of data collection (Barros 2006). NOAA National Centers for Environmental Information (NOAA 2016) provides rainfall data at the point scale, often at hourly intervals.

Gridded rainfall data are also available using two methods. The first approach, which produces rainfall grids through interpolation of point measurements, is based entirely on the assumption of spatial correlation of rainfall point measurements. Thus, the accuracy is dependent on the spatial density of rain gauges and how terrain influences the correlation of precipitation measurements. The second method, called data assimilation, uses weather models to infer spatial correlation and temporal evolution of rainfall, and then when combined with point measurements, adjusts model-predicted rainfall toward observed values.

Table 1. Selected engineering design standard applications and the requirements of their precipitation data inputs

Engineering application	Example Analysis	Requirements of rainfall data		
		Format	Spatial	Temporal
Water supply management	Reservoir routing simulation and optimization (e.g. STELLA)	Time series	Basin scale (e.g. 1000 km ²) to point scale (<1km ²)	Seasonal to daily
Water quality monitoring	Water quality analysis simulation (e.g. streamflow simulation)	Time series	Riverine scale (10 km ² or less)	Seasonal to hourly
Flood forecasting	Continuous hydrologic simulation (e.g. SWMM)	Time series	1 km ² or less	Sub-hourly
	Flood plain delineation (e.g. HECRAS)	IDF Curve	Point scale (< 1km ²)	Daily to hourly
Stormwater management planning	Peak discharge estimation (e.g. rational method, TR-55)	IDF Curve	Point scale (< 1 km ²)	Daily to sub-hourly
Wastewater collection and treatment planning	Wet weather flow estimation (e.g. Peaking factor)	Monthly maximum	Basin scale (e.g. 1000 km ²) to point scale (>1km ²)	Monthly

Table 2. Temporal and spatial characteristics of sources of re-analysis observations for North American domains

Data Source	Spatial scale	Time-step	Period of record	Website
NCEP North American Regional Reanalysis (NARR), NOAA Earth System Research Laboratory	32 km	3-hour, daily, monthly	Jan 1, 1979 to Dec 31, 2015	https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/north-american-regional-reanalysis-narr
Observation-reanalysis hybrid dataset, Princeton University	1 degree (~110 km) 1/2 degree (~55 km) 1/4 degree (~25 km)	3-hour, daily, and monthly	Jan 1949 to Dec 2010	http://hydrology.princeton.edu/data.pgfp.php
ERA-Interim Global Atmospheric Reanalysis, ECMWF	80 km	daily, monthly	1979 to present	http://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/
Parameter-elevation Regressions on Independent Slopes Model (PRISM), Oregon State University (USDA, NOAA)	4 km	daily	1895 to present	http://www.prism.oregonstate.edu
Variable Infiltration Capacity (VIC) Hydrologic Model, University of Washington	1/8 degree (~12km)	daily	1950 to 2000	http://www.esrl.noaa.gov/psd/people/ben.livneh/

By systematically merging numerous observations (available at different resolutions from rain gauges, satellites, or radar) with weather model output, data assimilation creates gridded precipitation data that is uniform and consistent with simulated weather conditions. These assimilation data sets are called “re-analysis” data. The quality of gridded data, especially relating to precipitation and extremes, is variable by location and time period, due to the changing combination of observation density and quality as well as model bias (Dee et al. 2011; Dee et al. 2016; Bosilovich et al. 2008; Sun and Barros 2014).

Reanalysis data are publicly available for multiple temporal and spatial resolutions for the North American domain (Table 2). Additional information can be found at the University Corporation for Atmospheric Research (UCAR) Climate Data Guide website (Dee et al. 2016) or at the reanalysis site maintained by the University of Colorado at Boulder (Reanalysis.org 2016).

Step 2. Access appropriate climate model output based on requirements for the existing application

Open source downscaled model output is becoming increasingly prevalent and diverse as data sources continue to emerge and climate models continue to evolve, yet model outputs are only publically

available at specific spatial and temporal resolutions that may not be consistent with the resolution of the inputs required for the engineering application or standard. Model outputs are also only provided for specific historical and future dates, and the length of this simulation period may not be equivalent to the length of rainfall record (e.g., 50 - 100 years) utilized in some methods to inform standards. These higher spatial-resolution (4–50 km) outputs are created using *downscaling models* that reduce the coarse resolution of the global atmospheric models (GCMs) (75–250 km) that are used as input (Cooney 2012; Di Luca et al. 2015; McGuffie and Henderson-Sellers 2001).

Characteristics of the sources of downscaled model output differ as a result of three main factors: (1) choices made by climate scientists in the downscaling process, including the downscaling method and the number of GCMs and emissions scenarios used; (2) consequences of the computational power and data storage that were available to the climate modelers, affecting length of simulation, and temporal and spatial resolution, which are aspects most relevant to engineers; and (3) decisions made by the climate modelers to store and allow access to the data. The next step in the framework provides context and information to allow an engineer to select and extract downscaled data, most suitable to the engineering application, from the myriad of available sources. Figure 2

presents a comparison of characteristics of six publically available sources of downscaled climate model output for North America, which include:

1. The North American Regional Climate Change Assessment Program (NARCCAP) [narccap.ucar.edu] (Mearns et al. 2008);
2. The North American Coordinated Regional Climate Downscaling Experiment (NA-CORDEX) [na-cordex.org] (Mearns et al. 2013);
3. The United States Geological Survey (USGS) regional climate model (RegCM3) [regclim.coas.oregonstate.edu] (Hostetler et al. 2011);
4. Eighth degree-CONUS (Continental U.S.) Statistical Asynchronous Regional Regression Daily Downscaled Climate Projections (ARRM) [cida.usgs.gov] (Stoner et al. 2013);
5. The Downscaled CMIP3 and CMIP5 Climate

and Hydrology Projections, from the Coupled Model Intercomparison Project (CMIP) versions 3 and 5 from the Bureau of Reclamation (and other collaborators) (Reclamation) [gdo-dcp.ucllnl.org] (Brekke et al. 2013); and

6. Multivariate Adapted Constructed Analogs (MACA) [maca.northwestknowledge.net] (Abatzoglou and Brown 2012).

As shown in Figure 2, model output from each of these sources varies depending on three main factors: Model Output Attributes, Model Simulation Choices, and Extraction and Access Features, which are discussed in subsequent subsections. Approaching the figure from top to bottom, a downscaled model source is selected based on desired characteristics.

	Model Simulation Choices						Model Output Attributes							Extraction and Access Features			
Source of output	Global climate model group		Emission scenarios		Nb. model simulations	Downscaling method	Spatial resolution		Historical period		Future period		Time step	Units	Date format	File format	Data access
	CMIP3	CMIP5	SRES	RCP	(does not include same model multiple runs)	Dynamical	4 km	6 km	1950	1958	2000	2015	3 hr	flux	calendar	csv	user interface
						Statistical	12 km	15 km	1965	1973	2030	2045	daily	(kg/m2/s)			
							25 km	50 km	1980	1988	2060	2075	monthly	depth			
NARCCAP	CMIP3				12	Dynamical			1970		2040	2070	3 hr	flux			
			A2					50 km		2000					360 dpy	netCDF	server download
NA-CORDEX		CMIP5		4.5	12	Dynamical			1950		2000		3 hr	flux			
				8.5			25 km	50 km		2000		2099	daily		360 dpy	netCDF	server download
					3	Dynamical		15 km				2020	monthly	flux			user interface
RegCM3	CMIP3								1960				decadal		365 dpy	netCDF	
			A2						2000		2099						
ARRM	CMIP3		B1				12 km		1960		2000		daily	flux	calendar	csv	
			A1									2099		depth		netCDF	server download
			A2		16	Statistical			2000								
Reclamation	CMIP3	CMIP5	B1	2.6			12 km		1950		2000		daily	flux	calendar	csv	user interface
			A1B	6.0		Statistical						2099	monthly	depth		netCDF	
			A2	8.5	20+				2000								
MACA		CMIP5		4.5			4 km	6 km	1950		2000		daily	flux	calendar	csv	user interface
				8.5	20+	Statistical			2000		2099			depth		netCDF	

Figure 2. Comparison of characteristics of six publicly available sources of downscaled climate model output for North America.

Approaching from left to right, the user views available choices, attributes, and features for a particular data source. A climate modeler begins

the downscaling process from far left column (selection of the global climate model group) then advances to the far right (providing access to

output). An engineer often begins the process of using climate model output at the far right (attempting to access the output) and subsequently makes choices from right to left.

Model Output Attributes

The process of selecting an appropriate model source should begin with the characteristics of the model output that are most relevant to the engineering application, such as the time step, or size of the grid cell (see Figure 2). For many water resources applications, the most limiting of these characteristics is the temporal resolution. Daily data is available from all sources; however, only Regional Climate Models, such as models in NARCCAP and NA-CORDEX projects, are able to produce rainfall output on a sub-daily level of 3-hours or less. When the engineering application is not limited to a sub-daily time step, additional resources are available at the daily level through sources that utilize “empirical downscaling” techniques, which rely on existing statistical relationships between large-scale climate systems and local weather patterns (Abatzoglou and Brown 2012; Khan et al. 2006; Murphy 1999; Chen et al. 2013). These techniques are less computationally intense than Regional Climate Models (RCMs), and can provide higher resolution output (4 - 12 km) for long simulation periods (1950 - 2100), and multiple emissions scenarios and global climate models (Cooney 2012).

RCMs are able to provide a finer temporal resolution because they use a technique called “dynamical downscaling” that provides physical characterization of weather processes occurring on small scale and contributing to precipitation (Anderson et al. 2003; Xu 1999; Musau et al. 2013). These models require large amounts of computational power and thus only limited scenarios (e.g., emissions) can be computed. For example, a 30-year, 50-km resolution simulation of the Water Resources Foundation (WRF) regional model on a supercomputer (with 240 processors) lasts for over 2 days; and a 150 year, 25 km simulation lasts for 90 days (Mearns et al. 2013). Prior to the release of NA-CORDEX in 2017, the requirement for a sub-daily time step restricted the user to NARCCAP data, a pioneering project that in 2006 compiled consistent output from numerous regional climate modelers across the globe.

Modeling scenarios were limited to a single emission scenario (SRES A2), a relatively low resolution (50 km), and short simulation periods (30 years); meaning the end-user did not have flexibility to select different characteristics. NA-CORDEX will provide longer simulation periods (1950 - 2100), a higher spatial resolution (25-km) and two emissions scenarios (RCP 4.5 and 8.5), providing more flexibility.

Model Simulation Choices

At the daily level, the user now has more flexibility to select a downscaled data source based on several data characteristics, including others that are relevant to engineering applications, like spatial resolution; as well as those in the first group in Figure 2, which are a result of model simulation choices made by climate scientists in the downscaling process: *global climate model group*, *emissions scenarios*, and *downscaling technique*.

Global climate model group refers to the version of the CMIP that was used to evaluate the GCMs used for downscaling. Currently in its sixth phase (Eyring et al. 2016), the CMIP was established in 1995 as a standard experimental protocol to compare GCM outputs. Downscaled output discussed in this manuscript originated from global climate models from either: (i) the CMIP3, which used early generation models and many evaluations have been completed; or the (ii) CMIP5, which used more experimental GCMs, and has a shorter record of development and evaluation (Taylor et al. 2011). Comparison between CMIP3 and CMIP5 indicates minor differences in future projections (Reichler and Kim 2008), and a large majority of comparisons do not address engineering-specific metrics (Wuebbles et al. 2014). The Bureau of Reclamation dataset is the only source to provide output from both CMIP3 and CMIP5 models. A compelling reason to prefer one over the other for precipitation analysis has not been presented, which means engineers have flexibility in choosing from either data set or a combination thereof, depending on which is best suited for their application.

One may wish to select a downscaled data source based on the *downscaling technique* (dynamical or empirical); however, there is no consensus relating to which technique is superior, since both have advantages and disadvantages

(Prudhomme and Davies 2009; Fowler et al. 2007). Similar to the decision to choose between CMIP3 and 5, a single reason to prefer one downscaling technique to the other does not exist, and engineers should select a source suitable for the engineering problem rather than accessibility of a particular data archive.

Emission scenarios estimate the potential concentration of greenhouse gases (GHG) in the atmosphere, based on pathways of socio-economic, technological, and political factors. CMIP3 global models use emissions scenarios from the Special Report on Emissions Scenarios (SRES) (Nakicenovic et al. 2000) created for the Intergovernmental Panel on Climate Change (IPCC) 3rd Assessment report (Houghton et al. 2001); whereas, CMIP5 global models use emissions in the form of Representative Concentration Pathways (RCPs) (van Vuuren et al. 2011) created for the IPCC Fifth Assessment Report (IPCC AR5) (IPCC 2012). When it is available to choose between multiple emissions trajectories, the authors recommend analyzing at least two scenarios when possible: (a) an upper bound that will provide the most conservative estimate of future conditions for use in engineering practice, such as SRES A2 (projecting 2.0 - 5.1 °C of warming by 2100) or RCP 8.5 (5-6 °C by 2100), and (b) a lower bound that is aligned to targets of the Paris agreements (Framework Convention on Climate Change 2015), similar to SRES A1B and RCP 4.5. Choosing between emissions scenarios is most relevant for infrastructure of long lifetimes (40 years or more), since many impacts across emissions scenarios generally diverge after the middle of the century (Collins et al. 2013). Irrespective of the scenarios and model sources that are ultimately chosen by the engineer, it is important to document assumptions and make them available to those interpreting the findings.

Extraction and Access Features

Once a downscaled data source has been selected based on the desired characteristics (from Figure 2), the user should extract the output at the spatial grid closest to the geographical location of interest, for a historical simulation period (usually one for which the user has historical data from framework step 1) and a future simulation period (for dates and

length required for the engineering application). The user should obtain output from several climate models from within a data source in order to create an ensemble of downscaled model output. The initial ensemble should include all model simulations available for the selected emissions scenario(s), in order to accurately assess model performance and uncertainty, which is discussed in Step 3 of the framework.

To access the output, some sources provide a user interface, including the USGS, Bureau of Reclamation, and MACA. The USGS (RegCM3) and Bureau of Reclamation websites allow for selection of multiple grid cells and provide spatially averaged time series at different time steps. NARCCAP, NA-CORDEX, and the ARRM sources do not have the guidance of a user interface and data must be extracted as individual files from a server. These files are usually available in netCDF (network Common DataForm) format, and require software packages (available in Excel, MATLAB, R, python) to extract the time series of data at the desired geographical location(s). Precipitation stored in netCDF files is often provided as instantaneous flux values (in units of kg/m²s), which is converted to precipitation depth over a time period by dividing by the density of water (1,000 kg/m³) and multiplying by the number of seconds in the time step.

Step 3. Account for climate model uncertainty and reliability

Step 3 of the framework addresses the importance of analyzing model performance and examines the possibilities for bounding uncertainty of the group or ensemble of models selected in step 2. Downscaled climate models are susceptible to large uncertainties, which can be inherited by GCMs, or introduced in the downscaling process. The three types inherited from GCMs include: (i) scenario uncertainty of future GHG emissions, (ii) natural (internal) climate variability (initial conditions), and (iii) inter-model discrepancies (modeling assumptions) (Kirtman et al. 2013). Natural or internal variability for precipitation contributes the most uncertainty in the early 21st century, whereas inter-model uncertainty makes up the largest majority after 2040 (Kirtman et al. 2013; Hawkins and Sutton 2011). Scenario uncertainty is

managed through the use of multiple emissions scenarios, described in step 2. Internal conditions uncertainty is bounded by producing several simulations using different initial criteria (Knutti et al. 2009; Musau et al. 2013). Most sources of climate output, however, only provide output from a single simulation (or an average of multiple simulations) for each climate model, with the exception being the Bureau of Reclamation dataset.

Scientists recommend the use of multiple models, or a “model ensemble,” in order to avoid misleading conclusions from inter-model uncertainty (Barsugli et al. 2013), introduced from the GCMs or through the downscaling technique (Chen et al. 2011). However, which models to include in the ensemble will depend on the approach used to manage uncertainty, and possibly, the reliability of the climate models, discussed in the following subsections.

Climate model reliability

Climate models considered by scientists to be “more reliable” include those that are well documented, are well established (with many years to make improvements), and that produce stable results. Models may be considered less reliable if they produce output that is “biased” with respect to the observed metric(s). Bias (also referred to as model systematic error) is defined, in this context, as the average deviation between the observed value (or empirical statistic) and the values or statistics obtained from the historical climate model simulations. The deviation may be larger than zero for numerous reasons, including the assumptions and simplifications made in the modeling equations (of the global, regional, and/or statistical models). Bias can be overcome by changing internal modeling assumptions (although this often shifts bias in another direction) or through bias correction techniques applied to the simulated rainfall output. Bias is assessed through the comparison of observations to “hindcast” model simulations (i.e., simulation of historical conditions) at the spatial and temporal resolution of the climate model (Gleckler et al. 2008). While adequate performance in hindcasting is not a guarantee of reliability for future predictions, poor performance with historical conditions can be used to identify unrealistic models. For regional climate models, these hindcast runs are driven by historical

data, which is different from simulating the past using atmospheric conditions of GCMs. Hindcast runs are expected to be temporally consistent with historical data.

Nearly all models exhibit some instances of bias; however, the magnitude varies depending on the metric, season, and models examined (Hall 2014). NARCCAP regional climate models, which have not been bias-corrected, were able to estimate mean annual precipitation with relative precision (exhibiting low bias); however, extreme precipitation statistics (e.g., annual maxima, 20-year return period) were often overestimated. The Weather Research Forecasting Model (WRF), from NARCCAP, exhibited especially high bias for extremes, as the percentage error of the average maxima precipitation and the 20-year return value was greater than 90% for nearly all seasons and US regions (Wehner 2013). A study that examined 3-hour precipitation totals found that the Iowa State model (MM5) and the Scripps model (ECP) outperformed the Hadley model (HIRHAM), the Regional Climate Model (RegCM2), and the Canadian model (CRCM) (Anderson et al. 2003).

Statistical downscaling techniques usually account and correct for bias in the downscaling process; however, for a given metric, some techniques have been shown to outperform others. For extreme values, output from the ARRM statistical downscaling method, which includes bias-correction and cross-validation, showed improved accuracy and ability to be efficient and generalizable across regions (Stoner et al. 2013). For precipitation (and other variables), the MACA method has been found to outperform the Bias-Corrected Spatial Disaggregation method (BCSD), used to create the Bureau of Reclamation dataset, due to the ability to jointly downscale certain variables (Abatzoglou and Brown 2012).

Bounding Uncertainty

The approaches used to manage uncertainty are: the extremes approach, the ensemble approach, and the validation approach. The extremes approach examines the full range of future scenarios by using all output extracted in step 2; however, drawbacks include the time and effort expended to consider all models; and the possibility that the full range of models may produce an unrealistic representation, since some models may be unreliable (Fowler et al.

2007). The ensemble approach uses a weighted average of the climate model ensemble to develop a probability distribution of the range. However, there is no consensus on the minimum and maximum number of models to consider in this ensemble (Mote et al. 2011). One study found that skill converged after six or more models were included (Pierce et al. 2009). It is possible to weight models equally or based on reliability criteria. Wehner (2013) found that averaging the ensemble using complicated weighting schemes was not more effective than simply removing the unreliable or poorly performing models.

The validation approach is the final approach to managing uncertainty and involves “culling” the ensemble by removing unrealistic models based on performance criteria (Charles et al. 1999; Flato et al. 2013; Mote et al. 2011). Some studies have demonstrated that ranking the models based on performance leads to a difference in predictions (Gleckler et al., while others have shown that the differences due to model culling is slight (Mote et al. 2011); however, one study found that results for a precipitation metric were nearly indistinguishable between the average of the 11 best performing GCMs and 11 randomly selected GCMs, from the CMIP3 ensemble (Knutti et al. 2009). For engineers, however, culling may be an appropriate avenue in order to reduce ensemble size. While all uncertainty approaches provide utility, it must be highlighted that any estimation of uncertainty from a range of climate models will never provide perfect insight into the full spectrum of possible futures (Mote et al. 2011). The engineer should decide which approach is best suited to their application then clearly state all assumptions.

Step 4. Incorporate climate model output into the required engineering format

After the ensemble of desired climate models has been selected, the precipitation data projected by these models for the relevant future time frame must be incorporated into the existing method for developing the design standard (or application). Since climate model output is provided at a “gridded” resolution (4 km or higher), this step will often require adjustment of the model output to an even finer spatial resolution (e.g., $< 1 \text{ km}^2$). Model output may also need adjustment temporally, to

obtain a smaller time step. These adjustments are accomplished using additional downscaling or disaggregation techniques (Durrans et al. 1999) that depend upon the required format of the precipitation data needed to update the specific design standard. In order to adequately account for the range of uncertainty from the selected climate models, these further downscaling techniques should be applied individually to each model output before averaging and should not be applied to the average of the outputs, since this method filters data variation (Wehner 2013).

As discussed in step 1, sometimes the engineering application explicitly requires high-resolution time series (e.g., at the station scale, or at intervals smaller than 3 hourly) for analyses like streamflow simulation or flood forecasting. In this case, further statistical downscaling techniques must be applied to the ensemble that was selected in steps 2 and 3. Statistical downscaling methods include applying transfer functions, weather generators, weather typing, or quantile-mapping to the gridded, downscaled model output (Wood et al. 2004). Weather generators use empirical relationships calculated from observations to simulate synthetic time series for rainfall data (Andréasson et al. 2004; Chen et al. 2015). Weather typing, or resampling, involves relating the weather patterns of the larger scale climate model to observed patterns in the local area (Prudhomme et al. 2002; Onof and Arnbjerg-Nielsen 2009). Quantile mapping, also used for bias-correction, matches the empirical quantiles of re-gridded historical data to those of the historical climate simulation, then adjusts the future climate simulation based on the difference between the historical data and simulation (Boé et al. 2007; Laflamme et al. 2016; Gudmundsson et al. 2012).

If the engineering design application does not require a high-resolution time series, it may be possible to avoid using complex downscaling techniques. If the engineering method is instead based on a statistical analysis of the observed data (like IDF curves), statistics or methods may be altered instead of the rainfall timeseries. The change factor approach and the bias-correction approach have been employed to adjust statistical metrics. The change factor, or delta change, approach adjusts an observed statistic (usually at the point scale) to a future date using a ratio or

percentage that is calculated from the gridded, climate model output (Forsee and Ahmad 2011; Zhu 2012). Bias correction modifies the future, gridded value from the downscaled climate model based on the difference between the observed statistics (point scale) and past model simulation statistics (grid scale) or hindcast simulation statistics (grid-scale) (Arnbjerg-Nielsen et al. 2013; Boé et al. 2007; Chen et al. 2015; Wilks and Wilby 1999; Wood et al. 2000). Quantile mapping may be employed as a bias-correction technique (Boé et al. 2007). Detailed methods to accomplish this are described in the example section for depth-duration-frequency curves.

Step 5. Interpret results and incorporate changes into design practice

At this point in the framework, the user should have been able to incorporate trends from an ensemble of downscaled climate model outputs into an existing design standard or engineering application, producing a range of resulting scenarios. To make use of the range of results that incorporate future climate scenarios, current engineering practice must evolve to incorporate principles relating to uncertainty and risk. Uncertainty may be addressed using exhaustive or simplified approaches that build on the climate-model results. Robust Decision Making (RDM) is a technique based on principles of minimizing regret and achieving acceptable thresholds. RDM involves testing future designs against the full plausible range of futures obtained from step 4 (Groves and Lempert 2007; Hallegatte 2009; Lempert 2013; Espinet et al. 2015). Researchers are increasingly using RDM to address the challenge of uncertainty associated with climate change; however, these methods can be computationally intensive and may not be appropriate for all applications. Other approaches to addressing uncertainty include defining an acceptable risk level in order to select a design value or strategy from a range of possibilities (Karsten Arnbjerg-Nielsen 2011; Hallegatte 2009; Hallegatte 2014). When applying either method, decision-makers should favor strategies that are adaptable, reversible, or have no- or low-regret characteristics, meaning they provide benefits even if impacts of climate change are not as severe as projected (Hallegatte 2009; Olsen 2015).

In addition to uncertainty, engineering designs will also need to better incorporate principles of non-stationarity. This means addressing the fact that the infrastructure system is subject to one or more shifts in exogenous factors (climate, land-use, demand patterns) over the course of the operating lifetime (Kilgore et al. 2016). Best practices may include: (i) testing for non-stationary trends (before assuming them), (ii) explicitly defining the final year in the future that the structure is designed to operate to, with adequate performance, and (iii) defining what adequate performance means for each structure. The Mann-Kendal (MK) test can be used to detect a non-stationary trend in an underlying distribution (Cheng and AghaKouchak 2014; Katz 2013; DeGaetano 2009; Kilgore et al. 2016). If detected, non-stationarity can be addressed by expressing one or more parameters or variables as a function of time (Katz 2013); however, such factors must be calibrated and verified.

Application of Framework: Depth-Duration-Frequency Curves

The next section illustrates the framework as applied to rainfall duration frequency curves, using Pittsburgh, PA as a case study. The application focuses on updating of depth-duration-frequency (DDF) curves, which are a form of IDF curves that present rainfall as a depth (inches or mm), rather than an intensity (inches or mm per unit time). Each step describes the decisions made in order to update the curves to reflect future trends and uncertainties, following the framework.

Step 0. Define the existing design standard (or application) that relies on precipitation information

DDF curves provide estimates of the depth of rainfall that characterizes the potential for extreme storms to occur in a particular region. Storms are differentiated based on their duration and frequency, or probability, of occurrence. Duration refers to the length of time that precipitation occurs, and is selected by the engineer based on the length of the design storm (or time of concentration) used to calculate stormwater runoff for a specific method, e.g., the rational method, Technical Release-55 (TR-55). Frequency of

occurrence is described as either: (i) an exceedance probability, which is the probability that an event of specific duration and depth will be exceeded in one time period (often 1 year), or (ii) a return period, or recurrence interval, which is the inverse of the exceedance probability, defined as the average length of time between events of the same depth and duration (McCuen 2005). When the time period is equal to one year, the rainfall depth expected for a storm of 24-hour duration and 25-year return period is equivalent to the depth of precipitation over 24 hours that has a 4% chance of being exceeded in any year. The return period is selected by stakeholders based on the acceptable risk level for a design to fail or be inundated. Frequency curves for use in design standards are created regionally in the U.S. by NOAA, available from the NOAA Atlas 14, which consists of a compilation of precipitation frequency estimates for all U.S. states and political entities (Bonnin et al. 2006; PennDOT 2011). The north and southeast regions of the continental US have been recently updated (2015 and 2013, respectively); however, many western regions (e.g., Montana, Washington, Oregon) have not been updated since 1973 (Hydrometeorological Design Studies Center and NOAA's National Weather Service 2016). It is important to recognize, however, that significant challenges exist with these curves due in large part to the spatially sparse observed data used to cluster regions with similar characteristics of extreme rainfall (Barros 2006).

Step 1. Understand historical requirements for existing standard and retrieve data

DDF curves have historically been created based on the underlying distribution of extreme events that occur in long time series (50 to 100 years) of observed rainfall. This process is applied at different durations of rainfall (5 minutes to 72 hours) by aggregating the data to the appropriate interval before analysis. Two methods are used to extract the extreme events, also known as block maxima or tails, including: Annual Maximum Series (AMS), where the maximum event for each duration storm is extracted for each year of record, or Partial Duration Series (PDS), where all values are taken above a threshold (Kilgore et al. 2016; Bonnin et al. 2006; CSA 2012). The PDS method, also known as Peaks Over Threshold (POT) is able

to account for multiple extremes that may occur in a single year and is useful for short periods of record; however, thresholds may be difficult to select, and events within a year may not be hydro-meteorologically independent (Beguería 2005).

AMS data points are often fit to a Generalized Extreme Value (GEV) distribution, described by location, μ , scale, σ , and shape, ξ , parameters (Visser and Petersen 2012; Bonnin et al. 2006; Coles 2001; CSA 2012). The shape parameter (which can be greater than, less than, or equal to zero) determines the form of the distribution (e.g., Gumbel (Type I), Frechet (Type II) or Weibull (Type III)) (Coles 2001). GEV parameters may be estimated using maximum likelihood techniques (Katz 2013; CSA 2012). When using the AMS method, the rainfall depth for a given duration and return period, i.e., the recurrence interval depth (z_p), is found by relating the GEV parameters to the probability, as presented in Equation 1a and 1b.

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi} [1 - y_p^{-\xi}], & \text{for } \xi \neq 0 \\ \mu - \sigma \log y_p, & \text{for } \xi = 0 \end{cases} \quad \begin{matrix} [1a] \\ [1b] \end{matrix}$$

where $y_p = -\log(1 - p)$, p is the probability of exceedance in any year, and μ, σ , and ξ are the location, scale, and shape parameters of the GEV distribution, respectively (Coles 2001).

Step 2. Access appropriate climate model output based on requirements for the existing application

DDF curves are calculated for short duration (5 minutes to 12 hours) as well as long duration (24 to 72 hours) events. The statistically downscaled datasets (e.g., Bureau of Reclamation, ARRM, MACA) are suitable for long durations at the daily level or higher. However, the dynamically downscaled datasets (e.g., NARCCAP and NA-CORDEX) are more appropriate for this analysis, as they allow calculation of curves at shorter durations (e.g., the 3-hourly interval and greater). Sub-hourly durations would require additional temporal disaggregation or extrapolation techniques not undertaken in this demonstration.

NA-CORDEX output is recommended for use over NARCCAP, if available, since some models are available at finer temporal and spatial resolution (e.g., hourly, 25 km) and a longer

simulation period is available (1950 – 2100). However, at the time of this study, NA-CORDEX outputs were not yet available; thus, NARCCAP outputs were used. NARCCAP precipitation projections are produced using a single emissions scenario (SRES A2) at a 3-hour time step and spatial resolution of 50 km. Since only a single emissions scenario is available, this analysis does not account for scenario uncertainty; however, the A2 scenario is at the upper end of SRES scenarios and represents a conservative estimate of the future. Precipitation output were extracted for the single grid cell with the centroid nearest to the Pittsburgh International Airport (40.49° N, 80.24° W). Grid point maps are available to relate the geographical location of the grid cell to the associated (x,y) coordinates in the NetCDF data matrix. NARCCAP data were extracted for 11 available RCM-GCM simulations, which are regional climate model simulations (from six different RCMs) that use 1-2 global climate models as input. Simulations are available for historical (1970 – 2000) and future (2040 – 2070) periods (see Table 3). Data was also extracted for 6 hindcast runs, which are output from the RCMs after they were driven by historical reanalysis data (instead of a GCM). Time series were extracted for a single grid cell after downloaded data files (available in 5 year intervals for North America) were concatenated (Zender et al. 2016).

Step 3. Account for climate model uncertainty and reliability

The reliability of an ensemble of regional climate models can be assessed by comparing hindcasts of the regional models to historical observations. For NARCCAP, the re-analysis data are from the National Centers for Environmental Protection (NCEP) North American Regional Reanalysis (NARR) dataset (see Table 2). The reanalysis data act is input to the Regional Climate Model, and outputs from the Model, referred to as NCEP driven runs, are expected to reflect historical conditions. Reanalysis driven outputs are available on a 50-km resolution, 3-hour time step, for the time period from 1979 – 2006.

In this study, the reliability of the

NARCCAP RCMs was assessed by comparing the empirical distributions of the reanalysis outputs of the six regional climate models to those of observations obtained from the local stormwater authority in Pittsburgh (3 Rivers Wet Weather, 2015). Before comparison, the observed data, recorded on a 15-minute interval at 33 rain gauges throughout Allegheny County (area of 1,930 km²), was first scaled to the resolution of the reanalysis output (3-hour, 50-km) by aggregating to a 3-hour interval then averaging gauges within the 50-km grid cell.

The 3-hour exceedance probability, which represents the likelihood that a rainfall event of a specific volume will occur in a 3-hour period, was selected as the metric of comparison to represent the empirical distribution of both precipitation time series. The exceedance probability for each rainfall depth above zero (in the reanalysis and adjusted-observed time series) was calculated using a Weibull distribution, commonly used in precipitation analyses. Exceedance probabilities from the scaled-observations were plotted against those from the hindcast output (NCEP driven runs).

Uncertainty was bounded using the validation approach, which uses a performance or reliability analysis to select (or “cull”) models to include in the final ensemble. Three NARCCAP RCMs were selected, or culled, based on the visual proximity of the reanalysis exceedance curve to the adjusted-observed curve (Figure 3). The five RCM-GCM simulations available from the three selected RCMs were used in the subsequent analyses (Table 3).

Step 4. Incorporate climate model output into the required engineering format

After the performance assessment, data from downscaled models may be integrated into future DDF curves. Future trends may be incorporated in one of several steps taken to obtain the DDF values, including to: the underlying time series of the data record, the extreme value series (AMS or PDS), the GEV distribution, or directly to the return level intensities calculated from the distribution.

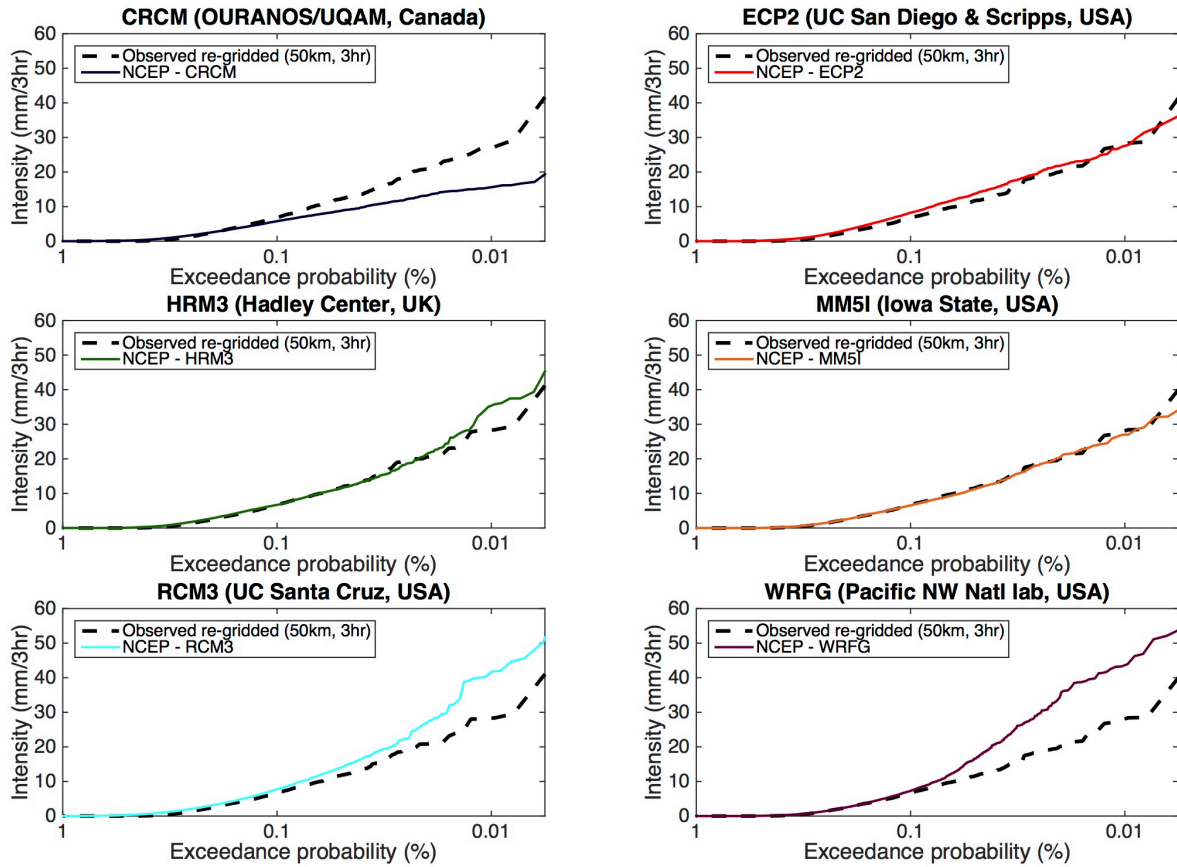


Fig 3. Comparison of 3-hr exceedance probabilities from 6 NCEP driven RCM runs in the NARCCAP ensemble (solid line) to observations re-gridded to 3-hr, 50km resolution (dashed line) for a grid cell in Pittsburgh region (1979–2014)

The first approach involves complex statistical downscaling techniques to obtain the appropriate temporal and spatial resolution of the time series. However, it has been hypothesized that if the engineer is only concerned with designing for extremes, it may be more manageable to avoid downscaling to a continuous time series, and instead adjust empirical quantiles through mapping functions (Hassanzadeh et al. 2013). A simple method that has been introduced in the engineering literature involves directly adjusting historical rainfall depths at the point scale, for a given return period and duration, based on the expected change from historical to future conditions at the grid scale (Zhu et al. 2012; Forsee and Ahmad 2011). Areal reduction factors have been employed to adjust the station scale rainfall, as reported by Zhu et al (2012), and summarized here in Equation 2.

$$I_F^{(s)}(T, d) = I_H^{(s)}(T, d) \frac{I_F^{(g)}(T, d)}{I_H^{(g)}(T, d)} \quad [2]$$

where I denotes the intensity for a given return period (T) and duration (d), at the station scale (s) or grid scale (g), for future (F) or historical (H) time periods.

In this analysis, climate signals are incorporated into regional DDF curves using areal reduction factors applied to historical depths at the station scale. This process has three stages: (1) historical DDF curves were recreated for the historical period available from the climate models (1970 – 2000) using airport station data (obtained from NOAA National Centers for Environmental Information); (2) change factors, or areal reduction factors, were calculated from DDF curves estimated from historical and future RCM gridded outputs; and (3) the change factors were applied to update historical curves. Steps (1) and (2) utilize

the same method for creating DDF curves, but on the native resolution of each data set. For the historical (1970 – 2000) and future (2040 – 2070) periods, return period depths values are calculated for the 3-, 6-, 12-, 24-, 48-, and 72-hour durations and the 2-, 5-, 10-, 25-, 50-, and 100-year return period. The moving window approach is applied to sum the underlying time series to the appropriate duration to obtain the annual maximum series. The AMS of each duration are fit to a GEV distribution using the method of moments and recurrence interval depths were calculated using Equation 1 for each 30 year period (1970 – 2000) and (2040 – 2070). GEV distributions are fit independently for the airport station data and each RCM. Change factors were calculated separately for each model as the ratio between the future and historical

gridded recurrence interval depths, and are applied to historical depths using Equation 2.

This simplified method is used solely for demonstration purposes of this framework. The method may be appropriate for understanding potential future trends in precipitation-frequency relationships; however, it is not a reliable alternative to more rigorous methods that alter the extreme value series or the GEV distribution parameters (Mailhot and Duchesne 2009; DeGaetano 2009; Cheng and AghaKouchak 2014; Shahabul Alam and Elshorbagy 2015). In the near future, NA-CORDEX will be available for a continuous time period (1950 – 2100) and could be used to inform a general trend in the future GEV distribution, notably the location parameter.

Table 3. Regional Climate Models and Associated GCM Drivers Compositing the NARCCAP Source of Downscaled Model Output; ECPC, and MM5I were Selected After Reliability Analysis (Step 3)

Model	Full Name	Modeling Group	GCM Driver
CRCM	Canadian Regional Climate Model	OURANOS / UQAM	CGCM3 (Third Generation Coupled Global Climate Model) CCSM (Community Climate System Model)
ECPC (ECP2)	Experimental Climate Prediction Center Regional Spectral Model	University of California-San Diego & Scripps Institute of Oceanography	GFDL (Geophysical Fluid Dynamics Laboratory GCM)
HRM3	Hadley Regional Model 3/Providing Regional Climates for Impact Studies	Hadley Centre	HADCM3 (Hadley Centre Coupled Model, version 3) GFDL (Geophysical Fluid Dynamics Laboratory GCM)
MM5I	MM5 – PSU/NCAR mesoscale model	Iowa State University	CCSM (Community Climate System Model) HADCM3 (Hadley Centre Coupled Model, version 3)
RCM3	Regional Climate Model version 3	UC Santa Cruz	GFDL (Geophysical Fluid Dynamics Laboratory GCM) CGCM3 (Third Generation Coupled Global Climate Model)
WRFP (WRFG)	Weather Research & Forecasting Model	Pacific Northwest National Lab	CCSM (Community Climate System Model) CGCM3 (Third Generation Coupled Global Climate Model)

Step 5. Interpret results and incorporate changes into design practice

Non-stationary conditions imply that the return period of an event will change with time. Mailhot and Duchesne (2009) state that design criteria under non-stationary conditions should explicitly consider (1) the expected lifetime of the structure, (2) that the probability of exceeding the design capacity and risk threshold will change over time, and (3) a statistical model that describes the

expected evolution of intense rainfall over time. The latter comes from the previous steps outlined in this framework and will include bounds of uncertainty represented as a range of plausible values for a given return period and duration. It is the responsibility of regulating agencies to provide guidance on which design value to choose within the range. Traditionally, design criteria have focused on selecting values as close to the expected value as possible, i.e. the mean of the range, assuming a normal distribution. One study

suggested that design levels should be selected as the higher-than-median-percentile of the design criteria in question (Karsten Arnbjerg-Nielsen 2011). Some argue that it is not possible to characterize the distribution and confidence intervals of the uncertainty, and an appropriate value cannot be selected independent of the decision being made (Hallegatte 2014). The authors of this study propose that instead of providing a single suggested value (and associated confidence intervals), agencies could provide two suggested values: a lower value, like the median of the range, and an upper value, like the 75th quantile, which could be used in either low or high-risk situations, respectively.

To address considerations (1) and (2) regarding infrastructure lifetime and changing risk level, Mailhot and Duchesne (2009) propose that the design engineer will need to establish two criteria: (1) the critical return period, the return period that the structure is designed to withstand, and (2) the reference year, the year in the future when the critical return period is reached. Values of the critical return period and the reference year will set expectations for the period of time expected for over design (if reference year is closer to end of lifetime) or under design (if reference year is closer to year of initial operation).

Mailhot and Duchesne also suggest that more severe guidelines are needed for infrastructure with long expected lifetimes (e.g., higher critical return periods and longer reference years), since these structures could experience extreme shifts in climate towards the end of life, at which they are most vulnerable to failure due to age and degradation of materials. Furthermore, where uncertainty in projections is especially high, designers may choose to select a shorter reference year to allow for adaptations to be implemented once conditions become more apparent. The 2009 study and the present authors stress the importance of implementing recurring performance evaluations of the drainage system in order to expose evolving system vulnerabilities. Adaptation strategies over time will be required to maintain an acceptable service level.

Results and Discussion of Framework Application

The following section focuses on results from steps 3 through 5 of the applied framework for the Pittsburgh, PA case study. From Step 3, performance analysis of the NARRCAP regional climate models; from Step 4, application of the change-factor to existing depth-duration-frequency curves; and from Step 5, relevance for stormwater design inputs and risk.

Figure 3 presents results from the performance analysis of the RCMs. The analysis compared the 3-hr exceedance probabilities from each of the six hindcast RCMs (NCEP driven runs) (1979 – 2006) to the exceedance probabilities of aggregated observations for the Pittsburgh region (2004–2014). The dashed line represents the 3-hour exceedance probability for the aggregated observations and the solid line represents the 3-hour exceedance of the hindcast RCMs. Proximity of the solid line to the dashed line represents similarity in the underlying empirical distributions and thus a higher skill of the RCM to represent historical statistics. These results show that for southwestern Pennsylvania, the Hadley Centre RCM (HRM3), Iowa State University RCM (MM5I), and University of California (UC) San Diego/Scripps RCM (ECP2) performed best in comparison to the other models, since these curves more closely agree with the dashed line (observed data). The Canadian RCM (CRCM) underestimates precipitation volume after the 1% exceedance probability, whereas the RCMs from UC Santa Cruz (RCM3) and from the Pacific Northwest National Laboratory (WRF3) overestimated the 3-hour precipitation depth after the 0.5% exceedance probability. Based on these visual results, the RCMs from the Hadley Center (HRM3), Iowa State (MM5I), and Scripps (ECP2) were selected for use in the subsequent analyses. Future research should examine quantitative metrics for objective selection of climate models based on reliability.

Figure 4 presents results for the change factors that were developed from the gridded climate model output. Change factors (CFs), i.e., the ratio of the future rainfall depth to the historical depth, are presented for return periods of 2-, 5-, 10-, 25-, 50-, and 100-years as separate sub-plots, and durations of 3-, 6-, 12-, 24-, 48-, and 72-hours within each return period plot. For each duration, the range of change factors represents the

variations between change factors from five different RCM-GCM simulations (3 RCMs and 3 GCMs). Change factors greater than 1.0 represent

an increase in the rainfall depth in the future; less than 1.0 denotes a decrease.

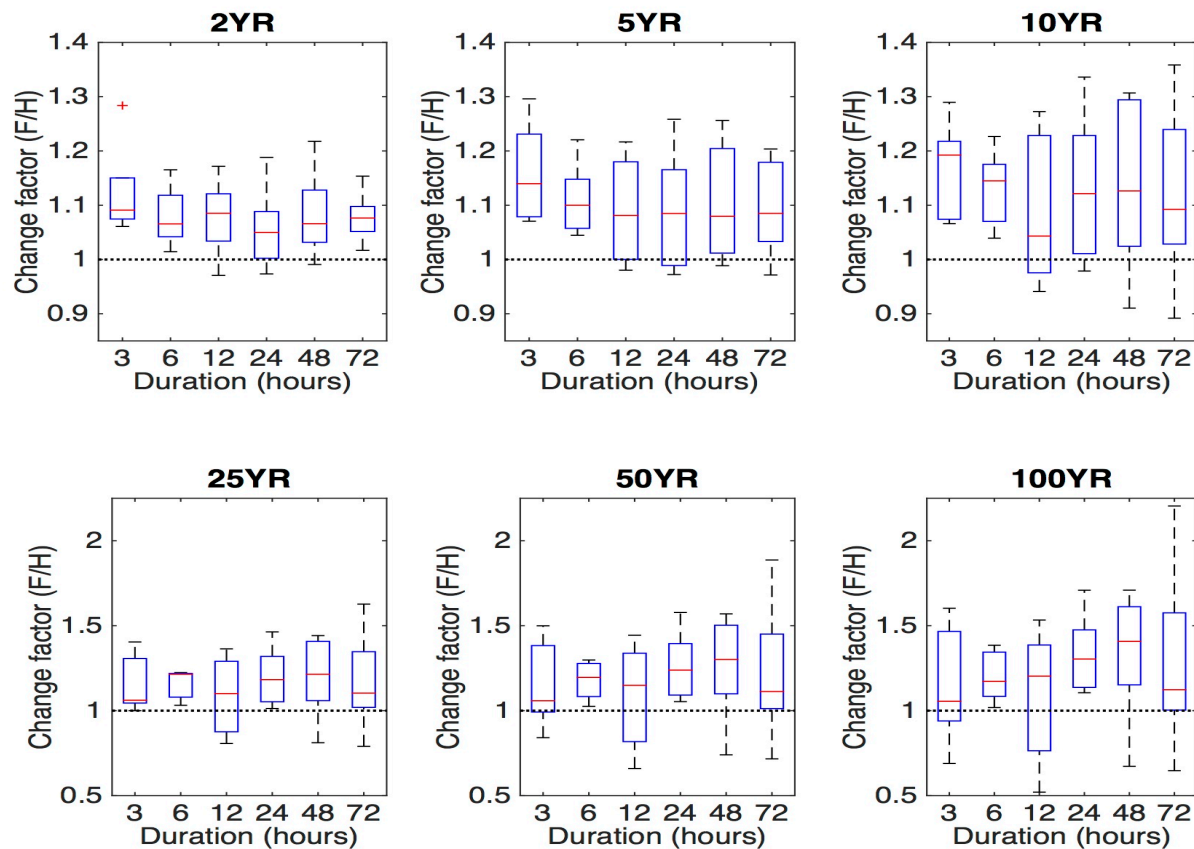


Figure 4. Range of change factors relating future (2040–2070) and historical (1970–2000) gridded rainfall depths for each model in the culled NARCCAP ensemble and a single grid cell in Pittsburgh; the middle bar in the box represents the median of the models, the top and bottom of the box plot represent the 25th and 75th quantiles, the whiskers extend to the 90th quantiles, and values outside these ranges are represented as plus signs.

The range of output from the 5 RCM simulations is presented as a box plot, where the bar in the box represents the median; the top and bottom of the box plot represent the 25th and 75th quantiles; the whiskers extend to the 90th quantiles; and plus signs represent values outside of the 90th quantile.

The median change factor for each duration and return period is larger than 1.0, which suggests that the depth of extreme precipitation is expected to increase in the future for Pittsburgh. With the exception of the 3- and 72-hour durations, the median change factor tends to increase as the return period increases. This is also the case for the 75th and 90th quantile change factor for all durations. This finding implies that the larger recurrence

interval storms (e.g., 25-, 50-, 100-year), for the same duration, may increase in severity at a sharper rate than the more frequently occurring storms (e.g., 2-year). It is also interesting to note that for the 2-, 5-, and 10-year return periods, the median change factor of the 3-hour duration storm is the largest of all durations. This is in line with other studies that found only short duration storms are shown to have consistently higher intensities in the future (Kuo et al. 2015; Cheng and AghaKouchak 2014); however, it also suggests further analysis of relative change is needed to produce understanding of whether the result has a clear physical interpretation and is expected to be reliably predicted across downscaling procedures and regions.

It may be possible to interpret change factors as a potential “climate safety factor” that could be applied to existing, stationary, depth-duration-frequency values. Based on these findings, a safety factor of 1.3 would encompass the majority of model uncertainty for depths of smaller return periods (e.g., 2 to 10 years); however, a factor of 1.3 is no longer valid when uncertainty magnifies as the return period increases to 25 years and larger. Change factors for extreme precipitation will vary depending on the duration and return period of the event, as well as the climate model, region, and future year analyzed; thus, additional studies are needed to determine appropriate climate safety factors by region, duration, and return period. As an alternative to applying a safety factor to existing curves, the authors recommend using values from

updated, non-stationary, depth- or intensity-duration-frequency curves.

Figure 5 presents the range of rainfall depths expected for the future period (2040 – 2070) based on the change factor method, for the previously listed durations and return periods. Change factors (reported in Figure 4) as less than 1.0 were converted to 1.0 for this analysis based on recommendations from the Canadian Standards Association, which state that beneficial aspects of climate change that allow for a reduction in design capacity should be neglected due to the inherent risks and costs that could arise from under-design (CSA 2012). When all models agree on findings suggesting change factors less than 1.0, this assumption should be reconsidered. To exemplify how rainfall depth changes with respect to the

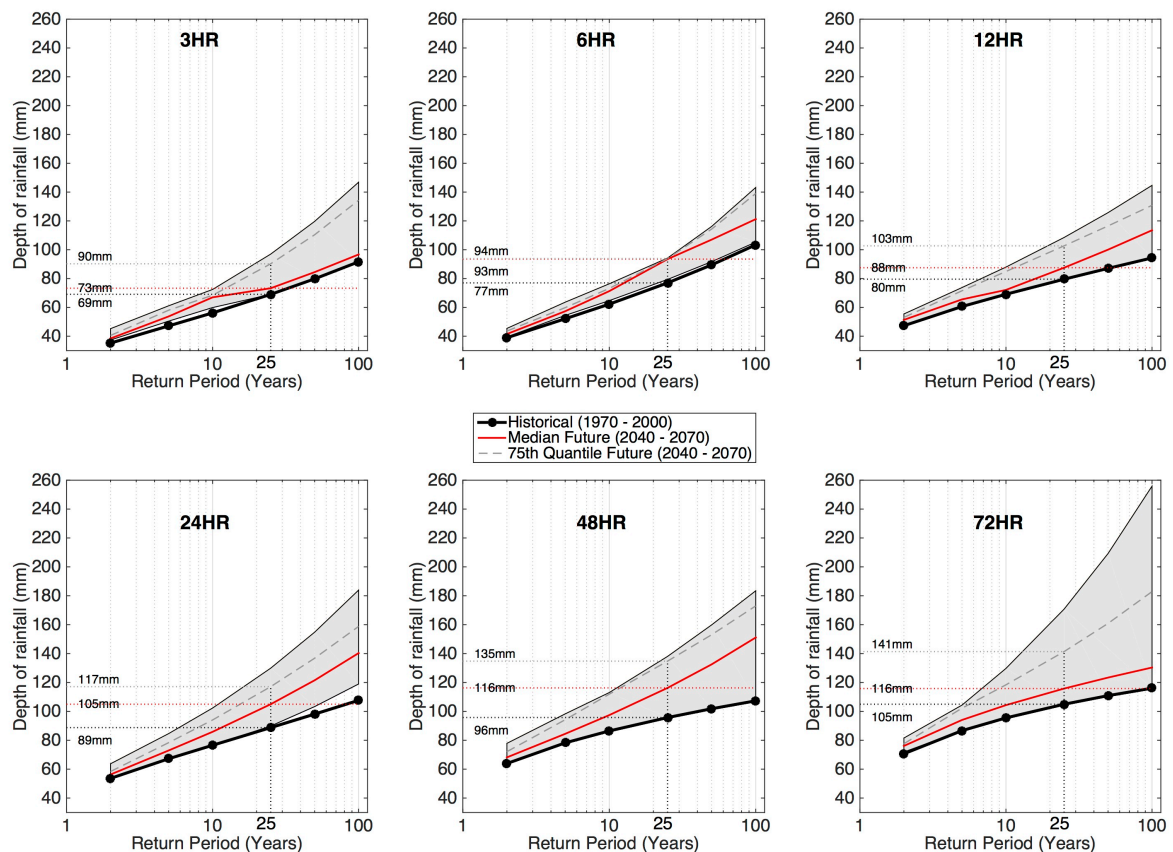


Figure 5. Updated DDF curves in Pittsburgh using change factor for the future period (2040–2070); the uncertainty is represented as the shaded grey area; the median is shown as the thin, dark, solid line; and the 75th quantile is shown as the thin, light, dashed line; the historical values (1970–2000) are shown as the thick, solid lines with markers; precipitation values in each plot, from bottom to top, respectively, represent the 25-year return period depth for the historical, future median, and future 75th quantile

probability of occurrence, results are portrayed for a specific duration, as a function of the return period. Uncertainty among the five models is represented as the grey region on the plot. The

Uncertainty of future projections tends to increase as return period increases. This phenomenon may be due to model variability of very extreme events; however, it is likely also a result of extrapolation of the GEV distribution to recurrence intervals larger than the underlying 30-year time series. One possible approach to overcome this limitation is to generate multiple simulations with the same model. Looking specifically at the 25-year return period, increases with respect to the historical depth are inconsistent across durations and do not increase monotonically as duration lengthens. The median, future, depth is equivalent to a 6%, 21%, and 10% increase for the 3-, 6-, and 12-hour durations, respectively, and an 18%, 21%, and a 10% increase for the 24-, 48-, and 72-hour durations, respectively. The 75th quantile depth ranges from a 21% to 41% increase from the historical depth, bounded by the 6-hour and 48-hour durations, respectively.

The future, median, 25-year depth can be extended horizontally right until it intersects with the historical line. This reflects the historical return period that would have been needed to ensure the 25-year return period performance in the future. For the 3-hour and 6-hour durations, this reflects the 35-year and 60-year return period, respectively; for 12-hour and 24-hour, it is about the 50-year and 85-year depths, and equal to or greater than the 100-year return period for durations 48 hours and larger. This finding indicates that merely doubling the return period (e.g., 25 year to 50 year) and using historical values may be appropriate for shorter duration storms (12-hours and less); however, this simplified method becomes inapplicable for larger duration storms. The historical 25-year depth (bottom, horizontal, dotted line) can also be extended left until it intersects the future, median curve (thin, solid line). The intersection suggests that designing for depths with respect to a stationary 25-year storm would only provide protection from the 7-to 12-year return period storms by 2070, for all durations.

These findings may be applied to the selection of the 25-year, 24-hour duration storm for

median of these models is shown as the thin, solid line, and the 75th quantile is the thin, dashed line. The historical values (1970 – 2000) are shown as the thick, dark line.

use as input to the TR-55 method, commonly used in storm water design for calculation of peak discharge. To do so, the authors assume the following: (i) that the updated curves represent the state of the art, (ii) the design structure is located on an arterial road of low traffic volume, and (iii) the reference year, the year after which infrastructure performance is not guaranteed, is 50 years. Assuming that the current year (2016) is equal to the year of conception of the project, the associated calendar year needed to describe the expected rainfall depth in the reference year is 2066, which falls within the future period evaluated in this analysis. The storm water structure represents a situation of low risk (due to placement on a low-volume arterial); thus, the authors recommend selection of the median depth of 105 mm for use as input to the TR-55 method.

Conclusions

This study presented a framework that may be used as a guide for agencies and engineers to update current infrastructure design standards to incorporate future, non-stationary trends. The framework begins by defining and understanding the existing requirements of the engineering application, and then discusses how to use this information to select and extract the most appropriate climate model source, and manage the associated model performance and uncertainty. The final steps examine options for adjusting model output for required temporal or spatial resolution of the existing engineering technique, and how to incorporate results into engineering practice by accounting for uncertainty, risk levels, and non-stationarity.

The general framework was applied to the updating of depth-duration-frequency (DDF) curves for a case study of Pittsburgh, Pennsylvania. Historical curves (recreated for the period from 1970 to 2000) were updated (for the future period of 2040 – 2070) using a change-factor approach. Change factors were developed using five historical and future regional climate model

simulations from NARRCAP, after climate model simulations were assessed for performance using historical rainfall data. The median change factor for each duration (3-, 6-, 12-, 24-, 48-, 72-hr) and return period (2-, 5-, 10-, 25-, 50-, 100-year) storm is larger than 1.0, which suggests that the depth of extreme precipitation is expected to increase in the future for Pittsburgh. Furthermore, a change-factor of 1.3 encompasses the majority of climate model uncertainty for rainfall depths of smaller return periods (2 to 10-years); however, this value may not be appropriate as uncertainty magnifies for lower frequency storms (25-year and above).

Results for the updated DDF curves indicate that merely doubling the return period (e.g., going from the 25-year to the 50-year frequency) and using historical curves may be appropriate for shorter duration storms (12-hours and less); however, this simplified method becomes inapplicable for larger duration storms. Similarly, results imply that designing for a rainfall depth equivalent to the future (2040-2070), median 25-year depth is comparable to designing for the historical (1970 – 2000) 50 to 100+ year depths, depending on the storm duration. If instead the designer selected a 25-year depth from the historical curve, this would be equivalent to the 7- to 12-year return period depths of the future, median value for various duration storms.

While future climate change is expected to introduce significant non-stationary changes and uncertainty, it is important to put these results into perspective of historical uncertainties. The authors used the years from 1970 to 2000 as the period of historical reference (since these dates coincided with the output available from the climate models); however, using this 30-year window as the sole basis for recreating historical curves introduces many challenges that were not addressed in this study. This historical period was assumed to be stationary; however, it is plausible that non-stationary trends in rainfall existed at this time (DeGaetano 2009). Non-stationary trends in historical data, as well as the spatial and temporal distribution of observations (e.g., (Barros et al. 2014)) could influence existing uncertainties associated with the empirical estimation of return period depths using GEV distributions, especially for return periods larger than the number of years of historical record. Given these factors, shifting or

extending the historical reference period has the potential to considerably alter the expected change in the future period. Thus historical reference periods and analyses should be selected with these factors in mind.

Recommendations and Future Work

Based on the results of the framework presented, and the ASCE initial guidance for adapting infrastructure and practice to a changing climate (Olsen 2015), the following information should be considered by engineers working with climate output for resiliency applications:

- Match intended engineering application with the appropriate climate model source;
- Different climate model sources require various amounts of effort for data extraction and preparation;
- Climate models have various levels of skill at representing historical mean and extreme statistical metrics and engineers need to understand the major issues and uncertainties involved;
- Create an ensemble and be transparent about assumptions;
- Test robustness of designs to extremes and alternative scenarios;
- Discuss tradeoffs and uncertainties in risk, resiliency, performance, and costs with stakeholders;
- Design for low-regret, adaptability, and robustness, and revisit designs when new information is available.

Because of stakeholder desires for enhanced resiliency to climate impacts, engineers will need to be familiar with choosing and incorporating climate change projections into planning and design. However, for engineering practitioners constrained by time and resources, it may not be feasible to expend the effort required for the detailed analyses described here. There is a need for collaboration across agencies and the research communities to serve as ad-hoc or standing boundary organizations that translate climate projections into relevant engineering information. Duties of these translational organizations may include providing rigorous standards for interpretation of climate data, understanding the utility of increasing the number

of models considered in an ensemble, development of a single, simplified user interface that accesses all downscaled data sources, and tools that automatically post-process data based on rigorous standards.

Acknowledgements

This research was supported in part by the National Science Foundation (NSF Collaborative Award Number CMMI 1635638/1635686), by the John and Claire Bertucci Fellowship, and by the Center for Engineering and Resilience for Climate Adaptation, in the Department of Civil and Environmental Engineering at Carnegie Mellon University. We are grateful to Dr. Jeanne VanBriesen for helpful comments, and to four anonymous reviewers whose suggestions greatly improved the manuscript. We acknowledge the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the WCRP's Working Group on Coupled Modelling (WGCM) for their roles in making available the WCRP CMIP3 multi-model dataset. Support of this dataset is provided by the Office of Science, U.S. Department of Energy. We also wish to thank the North American Regional Climate Change Assessment Program (NARCCAP) for providing the data used. NARCCAP is funded by the National Science Foundation (NSF), the U.S. Department of Energy (DOE), the National Oceanic and Atmospheric Administration (NOAA), and the U.S. Environmental Protection Agency Office of Research and Development (EPA).

References

- Abatzoglou, John T., and Timothy J. Brown. 2012. "A Comparison of Statistical Downscaling Methods Suited for Wildfire Applications." *International Journal of Climatology* 32 (5): 772–80. doi:10.1002/joc.2312.
- Anderson, Christopher J., Raymond W. Arritt, Zaitao Pan, Eugene S. Takle, William J. Gutowski, Francis O. Otieno, Renato da Silva, et al. 2003. "Hydrological Processes in Regional Climate Model Simulations of the Central United States Flood of June–July 1993." *Journal of Hydrometeorology* 4 (3): 584–98. doi:10.1175/1525-7541(2003)004<0584:HPIRCM>2.0.CO;2.
- Andréasson, Johan, Sten Bergström, Bengt Carlsson, L. Phil Graham, and Göran Lindström. 2004. "Hydrological Change – Climate Change Impact Simulations for Sweden." *AMBIO: A Journal of the Human Environment* 33 (4): 228–34. doi:10.1579/0044-7447-33.4.228.
- Arnbjerg-Nielsen, K., P. Willems, J. Olsson, S. Beecham, A. Pathirana, I. Bulow Gregersen, H. Madsen, and V.-T.-V. Nguyen. 2013. "Impacts of Climate Change on Rainfall Extremes and Urban Drainage Systems: A Review." *Water Science and Technology: A Journal of the International Association on Water Pollution Research* 68 (1). doi:10.2166/wst.2013.251.
- Arnbjerg-Nielsen, Karsten. 2011. "Past, Present, and Future Design of Urban Drainage Systems with Focus on Danish Experiences." *Water Science and Technology* 63 (3): 527–35.
- ASCE, and ASCE. 2013. Minimum Design Loads for Buildings and Other Structures. Standards. American Society of Civil Engineers. <http://dx.doi.org/10.1061/9780784412916>.
- Barros, A., and J. Evans. 1997. "Designing for Climate Variability." *Journal of Professional Issues in Engineering Education and Practice* 123 (2): 62–65. doi:10.1061/(ASCE)1052-3928(1997)123:2(62).
- Barros, Ana, Yajuan Duan, Julien Brun, and Miguel Medina. 2014. "Flood Nonstationarity in the Southeast and Mid-Atlantic Regions of the United States." *J. Hydrol. Eng.* 19 (10): 5014014. doi:10.1061/(ASCE)HE.1943-5584.0000955.
- Barros, Ana. 2006. "Rainfall Parameters for Design Storms Under a Changing Climate." In , 40. Washington D.C.: Center for Transportation and the Environment, North Carolina State University and USDOT.

- Barsugli, Joseph J., Galina Guentchev, Radley M. Horton, Andrew Wood, Linda O. Mearns, Xin-Zhong Liang, Julie A. Winkler, et al. 2013. "The Practitioner's Dilemma: How to Assess the Credibility of Downscaled Climate Projections." *Eos, Transactions American Geophysical Union* 94 (46): 424–25. doi:10.1002/2013EO460005.
- Beguéría, Santiago. 2005. "Uncertainties in Partial Duration Series Modelling of Extremes Related to the Choice of the Threshold Value." *Journal of Hydrology* 303 (1): 215–30.
- Boé, J., L. Terray, F. Habets, and E. Martin. 2007. "Statistical and Dynamical Downscaling of the Seine Basin Climate for Hydro-Meteorological Studies." *International Journal of Climatology* 27 (12): 1643–56.
- Bonnin, Geoffry M., Deborah Martin, Bingzhang Lin, Tye Parzybok, Michael Yekta, and David Riley. 2006. "NOAA Atlas 14: Precipitation-Frequency Atlas of the United States, Volume 2, Version 3.0." U.S. Dept Commerce. http://www.nws.noaa.gov/oh/hdsc/PF_documents/Atlas14_Volume2.pdf.
- Bosilovich, Michael G., Junye Chen, Franklin R. Robertson, and Robert F. Adler. 2008. "Evaluation of Global Precipitation in Reanalyses." *Journal of Applied Meteorology and Climatology* 47 (9): 2279–99. doi:10.1175/2008JAMC1921.1.
- Brekke, Levi, Bridget Thrasher, Edwin Maurer, and Tom Pruitt. 2013. "Downscaled CMIP3 and CMIP5 Climate Projections: Release of Downscaled CMIP5 Climate Projections, Comparison with Preceding Information, and Summary of User Needs." Technical Memorandum. Denver, CO: U.S. Department of the Interior, Bureau of Reclamation. http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/.
- Chandra, Rupa, Ujjwal Saha, and P.P. Mujumdar. 2015. "Model and Parameter Uncertainty in IDF Relationships under Climate Change." *Advances in Water Resources* 79 (May): 127–39. doi:10.1016/j.advwatres.2015.02.011.
- Charles, Stephen P., Bryson C. Bates, Peter H. Whetton, and James P. Hughes. 1999. "Validation of Downscaling Models for Changed Climate Conditions: Case Study of Southwestern Australia." *Climate Research* 12 (1): 1–14.
- Chen, Deliang, Christine Achberger, Tinghai Ou, Ulrika Postgård, Alexander Walther, and Yaoming Liao. 2015. "Projecting Future Local Precipitation and Its Extremes for Sweden." *Geografiska Annaler: Series A, Physical Geography* 97 (1): 25–39. doi:10.1111/geoa.12084.
- Chen, Jie, François P. Brissette, Diane Chaumont, and Marco Braun. 2013. "Performance and Uncertainty Evaluation of Empirical Downscaling Methods in Quantifying the Climate Change Impacts on Hydrology over Two North American River Basins." *Journal of Hydrology* 479 (0): 200–214. doi:10.1016/j.jhydrol.2012.11.062.
- Chen, Jie, François P. Brissette, and Robert Leconte. 2011. "Uncertainty of Downscaling Method in Quantifying the Impact of Climate Change on Hydrology." *Journal of Hydrology* 401 (3–4): 190–202. doi:10.1016/j.jhydrol.2011.02.020.
- Cheng, Linyin, and Amir AghaKouchak. 2014. "Nonstationary Precipitation Intensity-Duration-Frequency Curves for Infrastructure Design in a Changing Climate." *Scientific Reports* 4.
- Coles, Stuart. 2001. *An Introduction to Statistical Modeling of Extreme Values*. 1st ed. Springer Series in Statistics. London: Springer-Verlag London.
- Collins, M., R. Knutti, J. Arblaster, J.-L. Dufresne, T. Fichet, P. Friedlingstein, X. Gao, et al. 2013. "Long-Term Climate Change: Projections, Commitments and Irreversibility." In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley, 1029–1136. Cambridge, United Kingdom and New York, NY,

- USA: Cambridge University Press. www.climatechange2013.org.
- Cooney, Catherine M. 2012. "Downscaling Climate Models: Sharpening the Focus on Local-Level Changes." *Environmental Health Perspectives* 120 (1): a22–28. doi:10.1289/ehp.120-a22.
- CSA Standards. 2012. Development, Interpretation, and Use of Rainfall Intensity-Duration-Frequency (IDF) Information: Guideline for Canadian Water Resources Practitioners. Ontario, Canada: Canadian Standards Association.
- Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, et al. 2011. "The ERA-Interim Reanalysis: Configuration and Performance of the Data Assimilation System." *Quarterly Journal of the Royal Meteorological Society* 137 (656): 553–97. doi:10.1002/qj.828.
- Dee, Dick, John Fasullo, Dennis Shea, John Walsh, and National Center for Atmospheric Research Staff (Eds). 2016. "The Climate Data Guide: Atmospheric Reanalysis: Overview & Comparison Tables." January 21. <https://climatedataguide.ucar.edu/climate-data/atmospheric-reanalysis-overview-comparison-tables>.
- DeGaetano, Arthur T. 2009. "Time-Dependent Changes in Extreme-Precipitation Return-Period Amounts in the Continental United States." *Journal of Applied Meteorology and Climatology* 48 (10): 2086–99.
- Di Luca, Alejandro, Ramón de Elía, and René Laprise. 2015. "Challenges in the Quest for Added Value of Regional Climate Dynamical Downscaling." *Current Climate Change Reports* 1 (1): 10–21. doi:10.1007/s40641-015-0003-9.
- Durrans, S. Rocky, Steven J. Burian, Stephan J. Nix, Ahmed Hajji, Robert E. Pitt, Chi-Yuan Fan, and Richard Field. 1999. "POLYNOMIAL-BASED DISAGGREGATION OF HOURLY RAINFALL FOR CONTINUOUS HYDROLOGIC SIMULATION1." *JAWRA Journal of the American Water Resources Association* 35 (5): 1213–21. doi:10.1111/j.1752-1688.1999.tb04208.x.
- Erdman, Jon. 2016. "18 Major Flood Events Have Hit Texas, Louisiana, Oklahoma, Arkansas Since March 2015." *Severe Weather*. Weather Channel. August 16. <https://weather.com/storms/severe/news/flood-fatigue-2015-2016-texas-louisiana-oklahoma>.
- Espinet, X., A. Schweikert, and P. Chinowsky. 2015. "Robust Prioritization Framework for Transport Infrastructure Adaptation Investments under Uncertainty of Climate Change." *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, November, E4015001. doi:10.1061/AJRUA6.0000852.
- Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor. 2016. "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) Experimental Design and Organization." *Geosci. Model Dev.* 9 (5): 1937–58. doi:10.5194/gmd-9-1937-2016.
- Faturechi, R., and E. Miller-Hooks. 2014. "Measuring the Performance of Transportation Infrastructure Systems in Disasters: A Comprehensive Review." *Journal of Infrastructure Systems* 21 (1): 4014025. doi:10.1061/(ASCE)IS.1943-555X.0000212.
- Flato, Gregory, Jochem Marotzke, Babatunde Abiodun, Pascale Braconnot, Sin Chan Chou, William Collins, Peter Cox, et al. 2013. "Evaluation of Climate Models." In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by Isac Held, Andy Pitman, Serge Planton, and Zong-Ci Zhao. Cambridge, United Kingdom and New York, NY USA: Cambridge University Press.
- Forsee, W., and S. Ahmad. 2011. "Evaluating Urban Storm-Water Infrastructure Design in Response to Projected Climate Change." *Journal of Hydrologic Engineering* 16 (11): 865–73.

- doi:10.1061/(ASCE)HE.1943-5584.0000383.
- Fowler, H. J., S. Blenkinsop, and C. Tebaldi. 2007. "Linking Climate Change Modelling to Impacts Studies: Recent Advances in Downscaling Techniques for Hydrological Modelling." *International Journal of Climatology* 27 (12): 1547–78. doi:10.1002/joc.1556.
- Framework Convention on Climate Change. 2015. "Adoption of the Paris Agreement." United Nations. <https://unfccc.int/resource/docs/2015/cop21/eng/109r01.pdf>.
- Gleckler, P. J., K. E. Taylor, and C. Doutriaux. 2008. "Performance Metrics for Climate Models." *Journal of Geophysical Research: Atmospheres* 113 (D6): n/a-n/a. doi:10.1029/2007JD008972.
- Groves, D. G., & Lempert, R. J. (2007). A new analytic method for finding policy-relevant scenarios. *Global Environmental Change*, 17(1), 73-85. <https://doi.org/10.1016/j.gloenvcha.2006.11.006>
- Gudmundsson, L., J. B. Bremnes, J. E. Haugen, and T. Engen-Skaugen. 2012. "Technical Note: Downscaling RCM Precipitation to the Station Scale Using Statistical Transformations—a Comparison of Methods." *Hydrology and Earth System Sciences* 16 (9): 3383–90.
- Hall, Alex. 2014. "Projecting Regional Change Modeling Regional Climate Change: How Accurate Are Regional Projections of Climate Change Derived from Downscaling Global Climate Model Results?" *Science* 34 (6216): 1461–62.
- Hallegatte, Stéphane. 2009. "Strategies to Adapt to an Uncertain Climate Change." *Traditional Peoples and Climate Change* 19 (2): 240–47. doi:10.1016/j.gloenvcha.2008.12.003.
- Hallegatte, Stéphane. 2014. "Decision Making for Disaster Risk Management in a Changing Climate." In *Natural Disasters and Climate Change*, 177–194. Springer. http://link.springer.com/chapter/10.1007/978-3-319-08933-1_7.
- Hassanzadeh, E., A. Nazemi, and A. Elshorbagy. 2013. "Quantile-Based Downscaling of Precipitation Using Genetic Programming: Application to IDF Curves in Saskatoon." *Journal of Hydrologic Engineering* 19 (5): 943–55. doi:10.1061/(ASCE)HE.1943-5584.0000854.
- Hawkins, Ed, and Rowan Sutton. 2011. "The Potential to Narrow Uncertainty in Projections of Regional Precipitation Change." *Climate Dynamics* 37 (1): 407–18. doi:10.1007/s00382-010-0810-6.
- Hostetler, S. W., J. R. Alder, and A. M. Allan. 2011. "Dynamically Downscaled Climate Simulations over North America: Methods, Evaluation, and Supporting Documentation for Users." US Geological Survey. <http://regclim.coas.oregonstate.edu/>.
- Houghton, J.T., Y Ding, D.J. Griggs, M Noguer, P.J. van der Linden, X. Dai, K. Maskell, and C.A. Johnson. 2001. "Climate Change 2001: The Scientific Basis." Cambridge University Press, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, .
- Hydrometeorological Design Studies Center, and NOAA's National Weather Service. 2016. "Current NWS Precipitation Frequency (PF) Documents." US Department of Commerce. <http://www.nws.noaa.gov/oh/hdsc/current/pf.htm>.
- IPCC. 2012. "List of Major IPCC Reports." In *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC)*, edited by C. B. Field, V. Barros, T. F. Stocker, D. Qin, D. J. Dokken, K. L. Ebi, M. D. Mastrandrea, et al., 569–72. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.
- IPCC. 2014. *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Field, C.B., V.R. Barros, D.J.

- Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (Eds.]. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.
- Katz, Richard W. 2013. "Statistical Methods for Nonstationary Extremes." In *Extremes in a Changing Climate*, 15–37. Springer.
- Khan, Mohammad Sajjad, Paulin Coulibaly, and Yonas Dibike. 2006. "Uncertainty Analysis of Statistical Downscaling Methods." *Journal of Hydrology* 319 (1–4): 357–82. doi:10.1016/j.jhydrol.2005.06.035.
- Kilgore, Roger T., George Herrmann, Wilbert O. Thomas Jr., and David B. Thompson. 2016. "Hydraulic Engineering Circular No. 17, 2nd Edition Highways in the River Environment — Floodplains, Extreme Events, Risk, and Resilience." Federal Highway Administration.
- Kilsby, C.G., P.D. Jones, A. Burton, A.C. Ford, H.J. Fowler, C. Harpham, P. James, A. Smith, and R.L. Wilby. 2007. "A Daily Weather Generator for Use in Climate Change Studies." *Environmental Modelling & Software* 22 (12): 1705–19. doi:10.1016/j.envsoft.2007.02.005.
- Kirtman, B., S.B. Power, J.A. Adedoyin, G.J. Boer, R. Bojariu, I. Camilloni, F.J. Doblas-Reyes, et al. 2013. "Near-Term Climate Change: Projections and Predictability." In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley, 953–1028. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. www.climatechange2013.org.
- Knutti, Reto, Reinhard Furrer, Claudia Tebaldi, Jan Cermak, and Gerald A. Meehl. 2009. "Challenges in Combining Projections from Multiple Climate Models." *Journal of Climate* 23 (10): 2739–58. doi:10.1175/2009JCLI3361.1.
- Kuo, Chun-Chao, ThianYew Gan, and Mesgana Gizaw. 2015. "Potential Impact of Climate Change on Intensity Duration Frequency Curves of Central Alberta." *Climatic Change*, February, 1–15. doi:10.1007/s10584-015-1347-9.
- Laflamme, Eric M., Ernst Linder, and Yibin Pan. 2016. "Statistical Downscaling of Regional Climate Model Output to Achieve Projections of Precipitation Extremes." *Weather and Climate Extremes* 12 (June): 15–23. doi:10.1016/j.wace.2015.12.001.
- Lempert, Robert. 2013. "Scenarios That Illuminate Vulnerabilities and Robust Responses." *Climatic Change* 117 (4): 627–646. <http://link.springer.com/article/10.1007/s10584-012-0574-6>.
- Mailhot, A., and S. Duchesne. 2009. "Design Criteria of Urban Drainage Infrastructures under Climate Change." *Journal of Water Resources Planning and Management* 136 (2): 201–8. doi:10.1061/(ASCE)WR.1943-5452.0000023.
- Maurer, Edwin P., Levi Brekke, Tom Pruitt, and Philip B. Duffy. 2007. "Fine-Resolution Climate Projections Enhance Regional Climate Change Impact Studies." *Eos, Transactions American Geophysical Union* 88 (47): 504–504. doi:10.1029/2007EO470006.
- McCuen, Richard H. 2005. *Hydrologic Analysis and Design*. 3rd ed. Pearson Prentice Hall.
- McGuffie, K., and A. Henderson-Sellers. 2001. "Forty Years of Numerical Climate Modelling." *International Journal of Climatology* 21 (9): 1067–1109.
- Mearns, L. O., W. J. Gutowski, J. J. Barsugli, L. Buja, G. M. Garfin, D. P. Lettenmaier, and L. Leung. 2013. "The Development of North America CORDEX." AGU Fall Meeting Abstracts, December. <https://na-cordex.org>.
- Mearns, Linda, Ray Arritt, Dave Bader, Philip Duffy, William Gutowski, William Kuo, and Ruby Leung. 2008. "Collaborative Research The North American Regional

- Climate Change Assessment Program (NARCCAP): Using Multiple GCMs and RCMs to Simulate Future Climates and Their Uncertainty.” http://www.narccap.ucar.edu/doc/about/Project_summary.pdf.
- Milly, P. C. D., Julio Betancourt, Malin Falkenmark, Robert M. Hirsch, Zbigniew W. Kundzewicz, Dennis P. Lettenmaier, and Ronald J. Stouffer. 2008. “Stationarity Is Dead: Whither Water Management?” *Science* 319 (5863): 573–74. doi:10.1126/science.1151915.
- Mirhosseini, Golbahar, Puneet Srivastava, and Lydia Stefanova. 2013. “The Impact of Climate Change on Rainfall Intensity–Duration–Frequency (IDF) Curves in Alabama.” *Regional Environmental Change* 13 (1): 25–33. doi:10.1007/s10113-012-0375-5.
- Moss, R. H., G. A. Meehl, M. C. Lemos, J. B. Smith, J. R. Arnold, J. C. Arnett, D. Behar, et al. 2013. “Hell and High Water: Practice-Relevant Adaptation Science.” *Science* 342 (6159): 696–98. doi:10.1126/science.1239569.
- Mote, Philip, Levi Brekke, Philip B. Duffy, and Ed Maurer. 2011. “Guidelines for Constructing Climate Scenarios.” *Eos* 92 (31): 257–58.
- Murphy, J. 1999. “An Evaluation of Statistical and Dynamical Techniques for Downscaling Local Climate.” *Journal of Climate* 12 (8 PART 1): 2256–84. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-0033172619&partnerID=40&md5=fd6106d171b24b95b1b14fab06aa6e81>.
- Musau, J, J Sang, and J Gathanya. 2013. “General Circulation Models (GCMs) Downscaling Techniques and Uncertainty Modeling for Climate Change Impact Assessment.” In *Proceedings of 2013 Mechanical Engineering Conference on Sustainable Research and Innovation*. Vol. 5.
- Nakicenkoic, N, Joseph Alcamo, Gerald Davis, Bert de Vries, Joergen Fenhann, Stuart Gaffin, Kenneth Gregory, et al. 2000. *IPCC Special Report on Emissions Scenarios*. New York, NY and Cambridge, UK: Cambridge University Press.
- NOAA. 2016. “National Centers for Environmental Information.” Department of Commerce. <http://www.ncdc.noaa.gov/>.
- Olsen, J., ed. 2015. “Adapting Infrastructure and Civil Engineering Practice to a Changing Climate,” May. <http://dx.doi.org/10.1061/9780784479193>.
- Onof, C., and K. Arnbjerg-Nielsen. 2009. “Quantification of Anticipated Future Changes in High Resolution Design Rainfall for Urban Areas.” 7th International Workshop on Precipitation in Urban Areas 7th International Workshop on Precipitation in Urban Areas 92 (3): 350–63. doi:10.1016/j.atmosres.2009.01.014.
- PennDOT. 2011. “Chapter 7, Appendix A: Field Manual For Pennsylvania Design Rainfall Intensity Charts from NOAA Atlas 14 Version 3 Data.” In *PennDOT Drainage Manual 2010 Edition*, 1–40.
- Pierce, David W., Tim P. Barnett, Benjamin D. Santer, and Peter J. Gleckler. 2009. “Selecting Global Climate Models for Regional Climate Change Studies.” *Proceedings of the National Academy of Sciences* 106 (21): 8441–46.
- Prudhomme, Christel, and Helen Davies. 2009. “Assessing Uncertainties in Climate Change Impact Analyses on the River Flow Regimes in the UK. Part 1: Baseline Climate.” *Climatic Change* 93 (1): 177–95. doi:10.1007/s10584-008-9464-3.
- Prudhomme, Christel, Nick Reynard, and Sue Crooks. 2002. “Downscaling of Global Climate Models for Flood Frequency Analysis: Where Are We Now?” *HYDROLOGICAL PROCESSES* 16: 1137–50. doi:10.1002/hyp.1054.
- Reanalysis.org. 2016. “Advancing Reanalysis.” CIRES at University of Colorado Boulder. <http://reanalyses.org/>.
- Reichler, Thomas, and Junsu Kim. 2008. “How Well Do Coupled Models Simulate Today’s Climate?” *Bulletin of the American Meteorological Society* 89 (3): 303.

- Shahabul Alam, Md., and Amin Elshorbagy. 2015. "Quantification of the Climate Change-Induced Variations in Intensity–Duration–Frequency Curves in the Canadian Prairies." *Journal of Hydrology* 527 (August): 990–1005. doi:10.1016/j.jhydrol.2015.05.059.
- Stoner, Anne M. K., Katharine Hayhoe, Xiaohui Yang, and Donald J. Wuebbles. 2013. "An Asynchronous Regional Regression Model for Statistical Downscaling of Daily Climate Variables." *International Journal of Climatology* 33 (11): 2473–94. doi:10.1002/joc.3603.
- Sun, Xiaoming, and Ana P. Barros. 2014. "High Resolution Simulation of Tropical Storm Ivan (2004) in the Southern Appalachians: Role of Planetary Boundary-Layer Schemes and Cumulus Parametrization." *Quarterly Journal of the Royal Meteorological Society* 140 (683): 1847–65. doi:10.1002/qj.2255.
- Taylor, Karl E., Ronald J. Stouffer, and Gerald A. Meehl. 2011. "An Overview of CMIP5 and the Experiment Design." *Bulletin of the American Meteorological Society* 93 (4): 485–98. doi:10.1175/BAMS-D-11-00094.1.
- USEPA. 2015. EPA Storm Water Management Model (version 5.1). <http://www.epa.gov/water-research/storm-water-management-model-swmm>.
- Visser, H., and A. C. Petersen. 2012. "Inferences on Weather Extremes and Weather-Related Disasters: A Review of Statistical Methods." *Clim. Past* 8 (1): 265–86. doi:10.5194/cp-8-265-2012.
- Vuuren, Detlef P. van, Jae Edmonds, Mikiko Kainuma, Keywan Riahi, Allison Thomson, Kathy Hibbard, George C. Hurtt, et al. 2011. "The Representative Concentration Pathways: An Overview." *Climatic Change* 109 (1–2): 5–31. doi:10.1007/s10584-011-0148-z.
- Walsh, J.D., Wuebbles, K., and Kossin, J. 2014. "Ch. 2: Our Changing Climate. Climate Change Impacts in the United States." In *The Third National Climate Assessment*, 19–67. U.S. Global Change Research Program.
- Wastewater Committee of the Great Lakes - Upper Mississippi River. 2014. "Recommended Standards for Wastewater Facilities Policies for the Design, Review, and Approval of Plans and Specifications for Wastewater Collection and Treatment Facilities." Health Research Inc. <http://10statesstandards.com/wastewaterstandards.pdf>.
- Wehner, Michael F. 2013. "Very Extreme Seasonal Precipitation in the NARCCAP Ensemble: Model Performance and Projections." *Climate Dynamics* 40 (1–2): 59–80. doi:10.1007/s00382-012-1393-1.
- Wilks, D. S., and R. L. Wilby. 1999. "The Weather Generation Game: A Review of Stochastic Weather Models." *Progress in Physical Geography* 23 (3): 329–57. doi:10.1177/030913339902300302.
- Willems, Patrick, Jonas Olsson, Karsten Arnbjerg-Nielsen, Simon Beecham, Assela Pathirana, Ida Bülow Gregersen, Henrik Madsen, and Van-Thanh-Van Nguyen. 2013. "Climate Change Impacts on Rainfall Extremes and Urban Drainage: A State-of-the-Art Review." *Geophysical Research Abstracts* 15.
- Wood, A.W., L.R. Leung, V. Sridhar, and D.P. Lettenmaier. 2004. "Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs." *Climatic Change* 62 (1–3): 189–216. doi:10.1023/B:CLIM.0000013685.99609.9e.
- Wood, S. J., D. A. Jones, and R. J. Moore. 2000. "Static and Dynamic Calibration of Radar Data for Hydrological Use." *Hydrology and Earth System Sciences* 4: 545–54.
- Wuebbles, Donald, Gerald Meehl, Katharine Hayhoe, Thomas R. Karl, Kenneth Kunkel, Benjamin Santer, Michael Wehner, Brian Colle, Erich M. Fischer, and Rong Fu. 2014. "CMIP5 Climate Model Analyses: Climate Extremes in the United States." *Bulletin of the American Meteorological Society* 95 (4): 571–83.
- Xu, Chong-yu. 1999. "From GCMs to River Flow: A Review of Downscaling Methods and Hydrologic Modelling Approaches."

- Progress in Physical Geography 23 (2): 229–49.
<http://ppg.sagepub.com/content/23/2/229.abstract>.
- Zender, Charlie, Henry Butowsky, and Wenshan Wang. 2016. “Welcome to the netCDF Operator (NCO) Site.” source forge.
<http://nco.sourceforge.net/>.
- Zhu, J. 2012. “Impact of Climate Change on Extreme Rainfall across the United States.” *Journal of Hydrologic Engineering* 18 (10): 1301–9.
doi:10.1061/(ASCE)HE.1943-5584.0000725.
- Zhu, Jianting, Mark C. Stone, and William Forsee. 2012. “Analysis of Potential Impacts of Climate Change on Intensity–duration–frequency (IDF) Relationships for Six Regions in the United States.” *Journal of Water and Climate Change* 3 (3): 185–96.
doi:10.2166/wcc.2012.045.